

(19) JAPANESE PATENT OFFICE

(12) LAID-OPEN PATENT PUBLICATION (A)

(11) Publication number: HEI11-265388

(43) Date of laid-open publication: 28.9.1999

(51) Int. Cl.:

G 06 F 17/30

Numbers of claims: 5

Request of Examination: not yet requested

(21) Application number: HEI10-65284

(22) Date of filing: 16. 3.1998

(71) Applicant: NIPPON TELEGRAPH AND TELEPHONE <NTT>,
3-1, OOTEMACHI 2-CHOME, CHIYODA-KU, TOKYO

(72) Inventor: YASUHIRO AKIBA,
of NIPPON TELEGRAPH AND TELEPHONE <NTT>,
19-2, NISHISHINJUKU 3-CHOME, SHINJUKU-KU, TOKYO

(72) Inventor: ETSUO NAKAJI,
40-75 SAGANOARISUGAWACHO, UKYO-KU, KYOTO-SHI, KYOTO-FU

(72) Inventor: SHIGEO KANEDA,
13-6, OOEKITAFUKUNISHICHO 2-CHOME, NISHIKYO-KU, KYOTO-SHI, KYOTO-FU

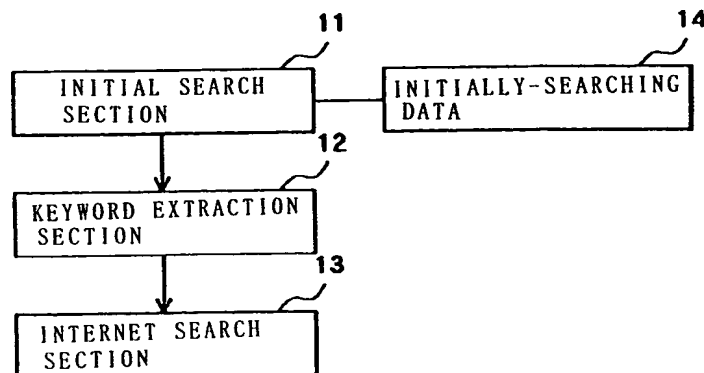
(74) Representative: PATENT ATTORNEY, TADASHI WAKABAYASHI, et al.

(54) Title of invention: Information Search Assisting Method and System, and Recording Medium
Storing an Information Search Assisting Program

(57) Brief Explanation of Disclosure

The disclosed technique seeks to allow a user, who has not yet clarified his or her request, to promptly identify the URL of a homepage favored by the user.

The user browses, by means of an initial search section 11, through initial-searching data (i.e., data to be initially searched through) 14 built in the hyper text format and existing near the user. Any screen fancied by the user is reported to a keyword extraction section 12, and the keyword extraction section 12 extracts a keyword from the reported user-fancied screen. Internet search section 13 transmits the keyword, extracted by the keyword extraction section 12, to a search server on the Internet so that information required by the user is searched for on the Internet. The initial-searching data 14 are prestored in a storage medium, such as a CD-ROM. This cited reference completely fails to disclose a pseudo screen similar in design and function to a screen of a music piece data setting site.



(19)日本国特許庁 (J P)

(12) 公 開 特 許 公 報 (A)

(11)特許出願公開番号

特開平11-265388

(43)公開日 平成11年(1999) 9月28日

(51)Int.Cl.⁸

G 0 6 F 17/30

識別記号

F I

G 0 6 F 15/40

15/419

3 1 0 F

3 2 0

審査請求 未請求 請求項の数5 O L (全 6 頁)

(21)出願番号

特願平10-65284

(22)出願日

平成10年(1998) 3月16日

(71)出願人 000004226

日本電信電話株式会社

東京都千代田区大手町二丁目3番1号

(72)発明者 秋葉 泰弘

東京都新宿区西新宿三丁目19番2号 日本
電信電話株式会社内

(72)発明者 中路 悦雄

京都府京都市右京区嵯峨野有栖川町40の75

(72)発明者 金田 重郎

京都府京都市西京区大枝北福西町2丁目13
番地6

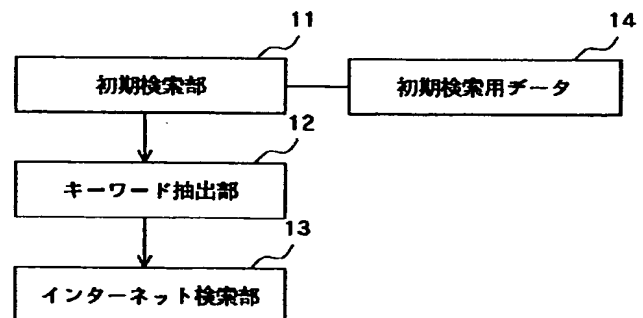
(74)代理人 弁理士 若林 忠 (外2名)

(54)【発明の名称】 情報検索支援方法、システムおよび情報検索支援プログラムを記録した記録媒体

(57)【要約】

【課題】 未だ自らの要求を明確化していないユーザが迅速に自分の気に入るホームページURLを明確化できるようにする。

【解決手段】 ユーザはハイパーテキスト形式により構成され、かつ、ユーザの近傍に存在する初期検索用データ14上を初期検索部11によりブラウジングする。ユーザが気に入った画面はキーワード抽出部12に報告され、キーワード抽出部12はユーザが気に入った画面からキーワードを抽出する。インターネット検索部13はキーワード抽出部12より抽出されたキーワードをインターネット上の検索サーバに送り込むことにより、ユーザが必要とする情報をインターネット上で検索する。



【特許請求の範囲】

【請求項1】 ハイパーテキスト形式により構成され、かつユーザの近傍に存在する初期検索用データ上をユーザが自由にサーフィンし、ある画面について興味深いことを発見すると、その画面が興味深いことを報告するステップと、

興味深いことを報告された画面からキーワードを抽出するステップと、

抽出されたキーワードをインターネット上の検索サーバに送り込み、ユーザが必要とする情報をインターネット上で検索するステップを有する情報検索支援方法。

【請求項2】 ハイパーテキスト形式により構成され、かつ、ユーザの近傍に存在する初期検索用データ上を、ユーザが自由にサーフィン可能な初期検索手段と、前記初期検索用データ中においてユーザが好ましいとして指定した画面中に含まれるキーワードを抽出するキーワード抽出手段と、

前記キーワード抽出手段により抽出されたキーワードをインターネット上の検索サーバに送り込むことにより、ユーザが必要とする情報をインターネット上で検索するインターネット検索手段を有する情報検索支援システム。

【請求項3】 前記初期検索用データがCD-ROM化されたハイパーテキスト形式の百科辞典であり、前記初期検索手段がWWWブラウザである請求項2記載の情報検索支援システム。

【請求項4】 前記キーワード抽出手段は、前記ブラウザの表示画面の一部をコピーしてキーワード保持ウィンドウに転記する請求項3記載の情報検索支援システム。

【請求項5】 ハイパーテキスト形式により構成され、かつ、ユーザの近傍に存在する初期検索用データ上を、ユーザが自由にサーフィン可能な初期検索処理と、前記初期検索用データ中においてユーザが好ましいとして指定した画面中に含まれるキーワードを抽出するキーワード抽出処理と、

前記キーワード抽出処理により抽出されたキーワードをインターネット上の検索サーバに送り込むことにより、ユーザが必要とする情報をインターネット上で検索するインターネット検索処理とをコンピュータに実行させる情報検索支援プログラムを記憶した記憶媒体。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】 本発明は、インターネット上で情報を検索するユーザを支援する情報検索支援方法およびシステムに関する。

【0002】

【従来の技術】 インターネットは、世界中に存在するホームページを渡り歩き（ネットサーフィンして）、ユーザの望む情報を自由に得ることができるシステムとして注目されている。しかし、現実には、人間の情報処理能

力は一定であるのに対して、インターネット上に設けられたアクセス可能なホームページは指数的にその個数を増加させており、爆発的に増加する情報量に対して、情報を受け取る側の情報選択能力が追いつかなくなっている。このような状況に鑑みて、ユーザが入力したキーワードを含むホームページ（その場所は、URLと呼ばれるアドレスにより識別される）を検索する、Yahoo、Goo、TITAN等の検索サーバが日本でも設置され、広く利用されている。

【0003】 検索サーバによる検索方法は2通りある。一つは、ユーザが思いついた検索用キーワードをタイプインして、それにより、当該キーワードを含むURLを検索してユーザに提示（複数解）する機能である。一方、適当なキーワードを思いつかないユーザのためには、「スポーツ」「パソコン」等のジャンル別に設けられたボタンがあり、例えば「スポーツ」を押下すると、さらにそのキーワードが細分された「サッカー」「野球」「ゴルフ」等のボタンが現れる。引き続いて、「ゴルフ」を押下すると、さらに細かい分類のキーワードがメニューとして現れる。この大まかな分類から細かい分類へと下りてゆく方法は、多段の段階を経て、ユーザの求めるホームページの一覧へとユーザをナビゲートしてゆくものであり、広く利用されている。

【0004】 また、自分の気に入ったホームページを探す方法として、自分が気に入った、あるいは、自分の興味のあるホームページを蓄積しておき、そこに現れるキーワードを自動的に抽出して、検索に利用する方法も多くの研究者により研究されている。この方法は、特定のジャンルや興味の対象を含むホームページを探す手法としては都合がよい。しかし、既に、一度、当該分野の「お気に入り」のホームページを探し当てた後でない、この手法は利用できない。

【0005】

【発明が解決しようとする課題】 ここでは、「お気に入り」のホームページが全く見つからない状況あるいは従来見つけたホームページとは異なるタイプの、全く新たなホームページを探すような要求をを考える。この場合、ユーザ自身も、最初は、何を探しているのかわかっていないことが多々あるであろう。そのような探査的なインターネットの利用を考える。この場合、過去に類似のホームページを探索していないので、現状の探索方法としては、適当に思いついたキーワードを入力してGoo等で探索するか、Yahoo等のトップダウンな検索を利用するしか方法がない。

【0006】 ここで、ユーザが「何か面白いものはないかな？」と思って、ユーザが検索サーバを利用したとする。この場合、「何か面白いものはないかな？」とキーワード入力することは勿論意味をなさないし、予め区分された概念的な分野キーワードである「スポーツ」「パソコン」の中に、それを見いだすことも（そもそも、何

が見たいかがわかっていないので)できない。すなわち、「インターネットで何か楽しもう」とするユーザにとっては、キーワードが予め既知でなければならない従来のインタフェース仕様は、決して快適なものではない。

【0007】このような「ふらふらと散策をしながら、自分の気に入ったものを探す」ためには、ハイパーテキスト形式が効果的であることはよく知られている。ハイパーテキスト形式では、とにかくある画面からスタートして、自分が興味をもった単語に連なるリンクをたどり、自分の要求を創発的に明確化しながら、望みの画面へと到達する。インターネットがこのハイパーテキスト形式を用いているのは、そのような、探索的な利用を狙いとしている。しかし、現実には、このような「そぞろ歩き」には、インターネットは不便である。何より、応答が遅いことと、ホームページの作者の好みにしたがった、かなり恣意的なリンクのみが張られているので、必ずしも、広いジャンルから、自分にふさわしい話題を自由に選べないからである。

【0008】では、トップダウンな分野選択手法により、探索サーバを利用しようとしても、これも、必ずしも利便性がよくない。一つは、ある程度の分野を降りていった先で、異なる分野に行きたいと思ったときには、また、分野指定を最初からやり直す必要があることである。もう一つの問題は、あくまで探索サーバ画面上のメニューを選択して分野を詳細化する点にある。このような、限られたメニュー名称から自分の希望するものを探しているのでは、ホームページの画面そのものを見て、興味を刺激されるハイパーテキスト上のサーフィンのような、創発的な検索ができないからである。

【0009】以上見てきたように、「何か面白いものはないかな?」と言った動機に基づくインターネットの利用には、既存の探索サーバを用いた検索、既存の「お気に入り」のホームページからのキーワード抽出による検索は、いずれも適当ではない。

【0010】本発明の目的は、上記従来技術の問題を解決し、未だ自らの要求を明確化していないユーザが迅速に自分の気に入るホームページURLを明確化できる情報検索支援方法およびシステムを提供することにある。

【0011】

【課題を解決するための手段】本発明の情報検索支援方法は、ハイパーテキスト形式により構成され、かつユーザの近傍に存在する初期検索用データ上をユーザが自由にサーフィンし、ある画面について興味深いことを発見すると、その画面が興味深いことを報告するステップと、興味深いことを報告された画面からキーワードを抽出するステップと、抽出されたキーワードをインターネット上の検索サーバに送り込み、ユーザが必要とする情報をインターネット上で検索するステップを有する。

【0012】また、本発明の情報検索支援システムは、

ハイパーテキスト形式により構成され、かつ、ユーザの近傍に存在する初期検索用データを、ユーザが自由にサーフィン可能な初期検索手段と、初期検索用データ中においてユーザが好ましいとして指定した画面中に含まれるキーワードを抽出するキーワード抽出手段と、キーワード抽出手段により抽出されたキーワードをインターネット上の検索サーバに送り込むことにより、ユーザが必要とする情報をインターネット上で検索するインターネット検索手段を有する。

【0013】ユーザは、身近の、しかし、世界のインターネットがもつ情報量に比べればデータ量的に小さいが、注意深く、すべてのジャンルをカバーするように記述されている初期検索用データ(例えば百科辞典)をサーフィンすることで、偏りのない検索を迅速に実行可能となる。また、最初は百科辞典をサーフィンしているので、その間には、インターネットにアクセスする必要がなく、従量制の課金でインターネットを利用しているユーザには、接続コストを下げるができる。

【0014】なお、従来、百科事典には最新の情報は入っていなかったが、近來のインターネットの発展により、インターネットで最新の項目を百科事典に追加するサービスが開始されようとしている。また、特に最新の流行や用語のみを集めた辞書(現代用語集)もあり、その様な、最新情報のみをサーフィンしたいユーザは、この様な最新情報のオンライン・アップデートや流行語のみのCD-ROMを利用すればよい。

【0015】

【発明の実施の形態】次に、本発明の実施の形態について図面を参照して説明する。

【0016】図1を参照すると、本発明の一実施形態の情報検索支援システムは、ハイパーテキスト形式により構成され、かつ、ユーザの近傍に存在する初期検索用データ14上を、ユーザが自由にサーフィン可能な初期検索部11と、初期検索用データ14中においてユーザが好ましいとして指定した画面中に含まれるキーワードを抽出するキーワード抽出部12と、キーワード抽出部12により抽出されたキーワードをインターネット上の検索サーバに送り込むことにより、ユーザが必要とする情報をインターネット上で検索するインターネット検索部13とから構成される。

【0017】ここで、ハイパーテキスト形式により構成された初期検索用データ14は、ハイパーリンクを辿って、ユーザが自由にブラウジングできるものであれば何でもよい。例えば、CD-ROM化されたハイパーテキスト形式の百科辞典がその例である。ただし、本発明の趣旨から言えば、ハイパーテキストを構成するインターネットのホームページは適当でない。インターネットでは、応答が遅く、また、特定のホームページからスタートしても、リンクが分野的に限定されているおそれがあるからである。この初期検索用データ14として、イン

ターネットサーフィンを用いることは、論理的には可能であるが、その応答時間の遅さから望ましくない。あくまで、ユーザが利用しているクライアント・マシン、あるいは、LAN等のデータ転送量が大きくてレスポンスの高速なネットワーク手段で接続されたサーバマシン上に設けられたハイパーテキスト形式のデータでなければならない。百科辞典は、注意深く、分野、用語を網羅的に掲載しているので、本発明の趣旨に最適である。

【0018】ユーザは、この初期検索用データ14上を初期検索部11により自由にブラウジングする（ステップ21）。すなわち、自由にサーフィンする。ハイパーテキスト形式により構成しているのは、ユーザがそれまでの検索履歴にとらわれず、思いつくままに、次々とリンクを辿れることを狙いとしている。それが、本来、ハイパーテキストの目的であるからである。そしてある画面について興味深いことを発見すると、その画面が興味深いことをキーワード抽出部12に報告する。この場合、ユーザの報告部としては、気に入ったことを示すボタンをユーザインタフェース上に設けておけば十分である。あるいは初期検索用データ14を構成する画面の中で、特定の画面にユーザが停留する時間が一定限度を超えたときに、「ユーザが当該画面を気に入った」と解釈してもよいこととなる。キーワード抽出部12は、ユーザが気に入った画面からキーワードを抽出する（ステップ22）。このキーワード抽出には、自然言語処理技術分野では既知のキーワード自動抽出手法（例えば、キーワードは名詞であることを提示して、名詞句の主名詞を選んだり、出現頻度の高い名詞を選ぶ等のアプローチが知れている）を用いればよい。また、そのような自動化手法を用いなくても、ユーザ自身にとって気に入った画面が現れる毎に、キーワードを、コピー・アンド・ペーストの形で、キーワード抽出部12に入力させるアプローチも考え得る。キーワードを収集した後、インターネット検索部13は、ユーザが必要とする情報をインターネット上で検索する（ステップ23）。この場合、インターネット検索部13は、Goo等の既存の検索サーバに対して、収集されたキーワードを渡して検索を実行するのみでよい。検索結果は、WWWへの応答として、ユーザに提示されるようにしておく。これにより、ユーザは、自分が気に入った分野、あるいはテーマを、インターネット上で発見することが可能となる。ユーザは、まず身近の、しかし、（世界のインターネットがもつ情報量に比べれば）データ量的に小さいが、注意深く、すべてのジャンルをカバーするように記述されている百科辞典をサーフィンすることで、偏りのない検索を迅速に実行可能となる。また、最初は百科辞典をサーフィンしているので、その間には、インターネットにアクセスする必要がなく、従量制の課金でインターネットを利用しているユーザには、接続コスト引き下げの効果がある。

【0019】図2は、初期検索部11とキーワード抽出

部12の例の、より詳細な説明図である。図2では、201はパソコンであり、Windows 95等のオペレーティングシステムにより稼動しているとする。CRTディスプレイ画面2011には、ここでは、WWWブラウザ2013と後述キーワード抽出部12を構成するところのウインドウ203が表示されている。また、ファイルシステム202は、搭載されているCD-ROMファイル2022の内容を、CD-ROMファイルアクセス部2021によりパソコン201に伝送する機能をもつ。ここで、CD-ROMファイル2022は、HTML形式のデータ表現をもつとする。HTMLについては、例えば、文献（ローラリメイ著「HTML入門」トッパン）を参照されたい。HTMLにより記述されたファイルは、WWWブラウザ2013により、ハイパーカード形式のデータとして表示され、自由にブラウジングできることは言うまでもない。

【0020】初期検索部11は、ここでは、WWWブラウザ2013により構成される。ユーザは、CD-ROMファイル2022の内容を自由にサーフィンして、その内容を検索することができる。

【0021】一方、キーワード抽出部12は、この開示例では、キーワード保存のために設けられたキーワード保持ウインドウ203により実現される。ユーザは、自分にとって、興味あるキーワードを見いだしたときには、パソコンのオペレーティングシステムが提供するコピー・アンド・ペースト機能（Windowsでは、CTRL+CとCTRL+Vの組み合わせとなる。）により、ハイパーカードの表示画面であるWWWブラウザの表示画面2012をコピーして、このキーワード保持ウインドウ203に転記することで、キーワードは抽出される。この例では、「相国寺」について、興味をもったので、キーワード保持ウインドウ203に転記した直後の状態を示している。

【0022】図3は、インターネット検索部13の動作ステップを、より詳細に説明したものである。処理を開始すると、図2のキーワード抽出部203により獲得されているキーワードは、Goo、TITAN、Yahoo等のインターネット上の検索サーバに転送される（ステップ31）。次に、検索がサーバ側で実行され（ステップ32）、検索結果はHTML形式のファイルとして、クライアント側に打ち返される（ステップ33）。これを、インターネット検索部13は受信する。受信された結果は、最終的に、ユーザに表示されるが（ステップ34）、HTML形式のファイルであるため、WWWブラウザを利用することが最も簡便である。

【0023】次に、実際に、この発明により、検索実験を行った結果を示し、その効果を定量的に説明しておく。開示された手法を、従来の手法（検索エンジンを利用したネットサーフィン）と実験的に比較した。評価項目、ならびに実験手法は、次の通りである。

【0024】評価項目

①検索所要時間

②得られた情報（見つけたHPから得た情報）に対する総合的満足度 5点満点

③得られた情報の質 5点満点

④得られた情報の量 5点満点

実験方法1：ネットサーフ法（本発明を利用する）

検索エンジンを利用して直接ネットサーフィン、リンク機能を利用する。検索エンジンは代表的なYahooとInfoseekを使用した。

【0025】実験方法2：辞典法

百科辞典ソフトを利用して、自分の趣向にあったターム

（キーワード）を抽出し検索エンジンにかける。百科辞典は、市販されている中から項目数の異なる以下の2種類から1個を選んで使用した。

【0026】

①マイペディア（平凡社）：項目数、約62,000

②エンカルタ（マイクロソフト）：項目数、約17,500

実験はネットサーフ法を20例、辞典法を24例実施した。実験結果を表1に示す。

【0027】

【表1】

	辞典法（平均値）	ネットサーフ法（平均値）
所要時間（分）	7.1667	8.45
得られたHPの質	3.875	3.2
得られたHPの量	3.875	3.2
総合的満足度	4.1667	3.2

これから以下がわかる。

【0028】・所要時間については、平均で差異が認められる。ただし、極端に大きな差がない。これは、目的としているホームページを見ている時間を含んで時間を計測しているためと、極めて高速のネットワークを利用しているためと思われる。64Kビット/秒以下の遅いネットワークでは、その差は、さらに顕著となると推定される。

【0029】・主観評価ではあるが、得られたHPの質や量については、統計的にも、明らかに差異が存在する。回線利用料金を必要としない状態で、キーワードを見だし得ること、そして、検索結果として、直接インターネットを利用するのと、同等以上の質が得られていることは、極めて興味深い。

【0030】なお、図1に示した各部の処理は情報検索支援プログラムとして、CD-ROM、FD、半導体メモリ等の記録媒体に記憶しておき、コンピュータにより読み出して実行するようにしてもよい。

【0031】

【発明の効果】以上のように、本発明によれば、ユーザが、「何を探したいか？」をわからないままに、検索を実行でき、全く新たな分野のホームページの発見が可能となる。

【0032】また、本発明によれば、直接にインターネットサーフィンによる探索に比べて、迅速に、サーフィンを行い、目的とするホームページに到着できる。

【0033】また、本発明によれば、当初の検索に百科辞典等の、偏りなく分野が網羅されたデータを利用できるため、迅速に目的とするホームページに到着できる。

また、偏りがないので、検索結果として得られるホームページの情報も、本人にとって、より満足度の高いものとなる。

【0034】また、本発明によれば、最初は百科辞典をサーフィンしているので、その間には、インターネットにアクセスする必要がなく、インターネット接続コストが削減される。

【図面の簡単な説明】

【図1】本発明の一実施形態の情報検索支援システムの構成図である。

【図2】図1のシステムの全体の処理の流れ図である。

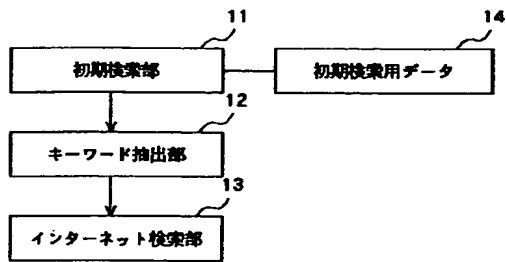
【図3】初期検索部11、キーワード抽出部12のより詳細な構成を示す図である。

【図4】インターネット検索部13の処理を示す流れ図である。

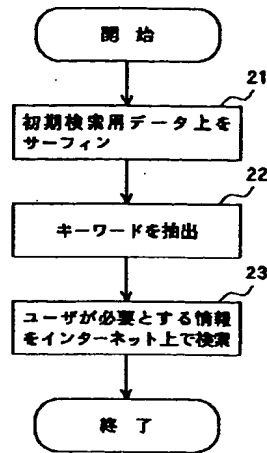
【符号の説明】

- 11 初期検索部
- 12 キーワード抽出部
- 13 インターネット検索部
- 14 初期検索用データ
- 21～23 ステップ
- 201 パソコン
- 202 ファイルシステム
- 203 キーワード検索ウインドウ
- 2011 CRT画面
- 2012 WWWブラウザの表示画面
- 2021 CD-ROMファイルアクセス部
- 2022 CD-ROMファイル
- 31～34 ステップ

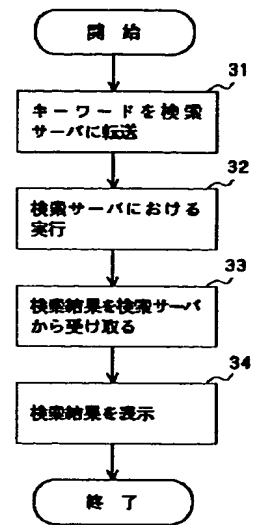
【図1】



【図2】



【図4】



【図3】

